

# CREATING, ANALYZING, AND USING A SMALL LEARNER CORPUS

**Charlotte Lin**

English Language Institute, Kanda University of International Studies

## ABSTRACT

*Learner corpora are made with "texts spoken or written by learners of the target language" (Cheng, 2012, pp. 171). Learner corpora are therefore informative and accurate profiles of learner language, which help teachers, as well as material and curriculum developers, to quickly and clearly evaluate learners' areas of strengths and weaknesses. Therefore, learner corpora can aid the development of short-term teaching plans, materials, long-term curricula and assessments that maximize learning outcomes for different target groups (O'Keeffe et al., 2007; Reppen, 2010; Flowerdew, 2012). Using a collection of short essays written by a learner of English, I will provide brief theoretical background information, explain the process and rationale of learner corpora construction, the specifications of the corpus design and construction, and analyze the corpus made by the text retrieval program WordSmith Tools. The analysis revealed the student's usage of present participles and gerunds in her writing, as well as the frequency and distribution of her errors, which helped improve future lessons. In addition, even though the study was based on the learner corpus of a French Canadian student, I will explain how to apply this method in the Kanda University of International Studies (KUIS) context.*

## INTRODUCTION

For this study, I applied various corpus linguistic techniques to analyze a collection of short essays written by a student whom I have been tutoring for one year in Quebec, Canada. Through the analysis, I aimed to map her usage of present participles and gerunds in her writing in order to identify the frequency of the distribution of her errors. The findings will be utilized to improve future lessons. In this paper, I will provide brief, theoretical background information of learner corpora construction and details of my learner corpus design and rationale. Additionally, there will be a discussion of the initial analysis made by the text retrieval program WordSmith. An extract of the corpus and the full result of the analysis are appended (Appendix A & B).

## CORPUS CONSTRUCTION & RATIONALE

A corpus is a large collection of authentic, written or spoken, computer-readable texts that have been compiled to serve a specific purpose. This purpose gives rise to the principles and

criteria of the corpus design and construction. Therefore, a corpus can be used for quantitative or qualitative language analysis (O’Keeffe *et al.*, 2007). The advent of using computers to study language has allowed for large amounts of data to be sorted, formatted, and analyzed with relative speed and ease, giving reliable results while minimizing human errors (McEnery *et al.*, 2006). Practical applications of corpus studies exist amongst fields such as professional communication (e.g. health care; law), forensic linguistics, lexicography, translation, and second language acquisition (e.g. testing; textbook design; learner corpora) (Flowerdew, 2012). I am particularly interested in the area of learner corpora because they are excellent tools for tracking learner progress and analyzing, for example, vocabulary acquisition and writing errors. They are also made with “texts spoken or written by learners of the target language” (Cheng, 2012, pp. 171). Learner corpora are therefore informative and accurate profiles of learner language, which help teachers and material and curriculum developers to quickly and clearly evaluate learners’ areas of strengths and weaknesses. They also contribute to the development of short-term teaching plans and materials, as well as long-term curricula and assessments that maximize learning outcomes for different target groups (O’Keeffe *et al.*, 2007; Reppen, 2010; Flowerdew, 2012).

The creation of learner corpora requires all the essential elements of corpora construction, such as representativeness, balanced sampling of authentic learner languages, and consideration of specific purposes in addition to the use of electronic texts. These elements address unique teaching problems and provide a rationale to the creation of specialized corpora as opposed to using available ones (McEnery *et al.*, 2006). Large-scale learner corpora, such as the *International Corpus of Learner English (ICLE)*, have led to the creation of learner-targeted materials (e.g. learner dictionaries) that benefit learners from numerous language and cultural background worldwide (Granger, 2003). However, even without access to these kinds of resources, educators can create their own learner corpora with available data from students. When creating such corpora, however, educators may only have access to small amounts of data. Flowerdew (2012) highlighted the following considerations for creating small and specialized corpora: 1) specific purposes for data compilation, 2) specific context (e.g. participant, communicative setting), 3) defined genres and discourse types, 4) subject area, and 5) variety of English.

According to these criteria, I have created a small learner corpus with one of my student’s weekly homework essays over a 10-week period. The specific purpose of my corpus is to analyze the usage of present participles and gerunds in her writings. The data (i.e. essays) is from a single participant, who is French-Canadian, in a business-English, one-to-one tutorial context. The genre of the articles is journal-style free writing about life events. The subject area is English as a Second Language (ESL) and the variety of English is learner English. I will explain the corpus design further in the next section, followed by a discussion of an initial analysis that I conducted with this corpus, including the analysis method, results and discussions.

## CORPUS DESIGN & CONSTRUCTION SPECIFICATION

The collection of the material came directly from the weekly writing assignment of my student, a French-Canadian woman from Quebec, over a period of 10 weeks. These articles were then compiled to create this learner corpus. The articles were written directly in Microsoft Word and sent to me via email. The materials, therefore, were already in electronic format. The articles were organized by week numbers (e.g. Homework 1 is the first article). In the analysis, only the raw versions of the articles were used. According to McEnery *et al.* (2006), no editing, extra-textual or linguistic annotations, such as parsing or error tagging, are needed in this type of analysis. Consequently, no editing was used in this study.

The written genre of these articles can be considered as free, personal, descriptive journal entries that recount stories or explore specific topics that the learner feels most strongly about at the time of writing. The entries vary across topics such as Valentine's Day, a dream, and political comments. This corpus can be considered specialized because it is genre-specific (McEnery *et al.*, 2006). Regarding the size, this corpus is constructed with 10 articles of roughly 500 words each, adding up to approximately 5000 words. The size of the corpus is small compared to any established corpora. However, regarding the issue of size, particularly in relation to its suitability, O'Keeffe, McCarthy & Carter state that the key "is often the design of a corpus as opposed to its size" (2007; pp. 4). In other words, the size should vary according to the design and purpose of the corpora. Reppen (2010) also emphasizes that for the purpose of language teaching, "small but representative is more important than millions of words," (pp. 55). Indeed, for the purpose of its creation, which is to evaluate the learner's progress over the 10-week period, this corpus offers a suitable representation of learner language and allows for an initial analysis of the errors within. The quality of representativeness, a key feature of any corpus, is achieved in this corpus by directly sampling materials from the learner's own authentic production of written texts. Additionally, the range of topics within this corpus adds variety in learner language, which is another characteristic of representativeness (McEnery *et al.*, 2006).

The corpus construction from the original Word files was straightforward. First, all 10 texts were re-saved as Plain Text files, and were subsequently marked up for WordSmith analysis. Only the "header" tag was used to mark the homework number and date (e.g. <header> HOMEWORK ONE: January 9th 2013 </header>). Tagging allows WordSmith to ignore these words in the analysis (Scott, 2013). The titles remained untagged because titles are considered part of language production in this analysis. The bodies of the articles remain untagged as well. A sample of corpus material, as shown in Appendix A, demonstrates how *Homework 9* is marked up in the corpus.

## ANALYSIS

In this section, I will provide rationale for the grammar topic chosen for the analysis, followed by the method of analysis, and finally, present the analysis results with discussion.

### Present Participles and Gerunds

Present participles, or words ending with *-ing*, are used in all progressive aspects of verbs (including present, past and perfect), in tenseless verb phrases and clauses to add extra information to main clauses, and as complements; they can also be used as gerunds, which are nouns that denote the action of a verb (Carter & McCarthy, 2006; Leech *et al.*, 2006; Parrott, 2010). The key difference between present participles and gerunds is the function, not form.

English present participles and gerunds present a unique challenge for French learners because of the absence of *-ing* verb forms in French (Walter, 2001). This causes French learners of English to make mistakes, often overusing *-ing* forms with state verbs, passive voice, adjectives, gerunds, and progressive aspects in all tenses (Walter, 2001; Parrott, 2010). My student has demonstrated a mix of correct and false usage in all these areas, and I used the following method to map her present participle and gerund usage in this corpus.

### Analysis Method

The method I used can be considered computer-aided error analysis because of the use of a text retrieval program called WordSmith to examine the frequency and patterns of present participle usage (Granger, 2003). *Concord* in WordSmith was used in this analysis to create a concordance from the corpus. *Concord* finds every instance of a node (i.e. keyword or phrase) in a corpus, and displays the search results in the Key Word in Context format, which places the node in the middle with their closest co-texts on both sides (Figure 1.) (Cheng, 2012). The nodes can be vocabulary or grammar related, and the resulting concordance allows for an overview and analysis of how the nodes are used in their respective contexts.

Due to the nature of present participles and gerunds (i.e. identical forms and a wide array of grammatical functions), the corpus analysis was straightforward yet resulted in generous insights into the learner language. By analyzing the corpus with the entry *\*ing* in *Concord*, WordSmith only searches for words containing *-ing* at the end and compiles all instances of *-ing* words into a concordance; in addition, by specifying the cutting of tags `<header>` and `</header>`, the tagged words were ignored by WordSmith in the analysis (Scott, 2013). I manually deleted lines which contained *-ing* words that were not verbs (e.g. *spring*, *during*, *something*, *the beginning*, *human being*, etc). As a result, there were 78 instances of *-ing* usage (Appendix B). I manually went through the list to identify and categorize correct and incorrect usages of present participles and gerunds. The next section discusses selected results in detail.

**Figure 1.** Screenshot of a part of the concordance list after “zapping” the non-verb –ing words, showing nodes (-ing words) in the middle of the concordance lines and immediate co-texts on both sides of the nodes.

N	Concordance	Set	Tag	Word #	Sen	Sen	Par	Par	lead	lead	Sec	Sec	File	%
1	a documentary on this new tendency of <b>living</b> in small apartments in Vancouver,			319	1950%		139%				054%		Homework 9.txt	63%
2	, I want only a «piéd à terre» and <b>finding</b> something else in the country. I			377	2233%		198%				076%		Homework 9.txt	76%
3	new «nest». I'll pass the next few days <b>finishing</b> me big clean-up of my actual			210	1153%		115%				042%		Homework 9.txt	43%
4	that will be too small. I try to figure out <b>living</b> in this «bubble». I know that I can			290	1770%		155%				058%		Homework 9.txt	58%
5	style. I have to find smart ways about <b>storing</b> things. This is my challenge for			439	2830%		250%				038%		Homework 9.txt	88%
6	and rapidly, because the market starts <b>decreasing</b> . I know that there are many			472	3030%		279%				035%		Homework 9.txt	95%
7	style with the doors which slide. I'm <b>looking</b> for a designer which is familiar			422	2725%		236%				035%		Homework 9.txt	85%
8	warm beaches of south; there are many <b>travelling</b> packages especially for this			173	877%		096%				035%		Homework 9.txt	35%
9	days which invite us being outside <b>practicing</b> Winter sports with friends			106	535%		059%				021%		Homework 9.txt	21%
10	weather is too grey, a lot people choose <b>being</b> on the warm beaches of south;			163	843%		031%				033%		Homework 9.txt	33%
11	of us. I love these days which invite us <b>being</b> outside practicing Winter sports			104	553%		058%				021%		Homework 9.txt	21%
12	for me is life. Ten year ago, I was <b>«moving»</b> with some people; dancer,			57	444%		030%				011%		Homework 8.txt	11%
13	<b>Moving</b> In my childhood, I always loved			1	0 7%		0 1%				0 0%		Homework 8.txt	2%
14	my body. I love to dance and to run and <b>moving</b> for me is life. Ten year ago, I			47	371%		024%				0 9%		Homework 8.txt	10%
15	will transform. This exercise combine <b>feeling</b> , mind and body. It's very			459	2533%		236%				037%		Homework 8.txt	87%
16	it's not as well that if I had a group who <b>witnessing</b> me in a studio. However,			496	2934%		353%				034%		Homework 8.txt	94%
17	, I still was on the campus McGill; I was <b>trying</b> to find my way toward my family			117	1048%		128%				023%		Homework 7.txt	23%
18	was younger that in reality, she was <b>preparing</b> a dessert for me. There also			150	1133%		136%				029%		Homework 7.txt	29%
19	that they were very well dressed, arrived <b>bringing</b> a large tray. It's at this time			233	1833%		156%				045%		Homework 7.txt	44%
20	I saw a horse, without cavalier, it was <b>bringing</b> a saddle, an English saddle.			35	439%		1 8%				0 7%		Homework 7.txt	8%
21	me and it lay down close to me, it was <b>protecting</b> me from the bad weather.			99	975%		123%				019%		Homework 7.txt	20%
22	Christmas and my mother who was <b>looking</b> younger? I really don't know.			286	2432%		139%				055%		Homework 7.txt	53%
23	by some specific concerns and I'm <b>cutting</b> off them for many raisons: I			450	3534%		244%				037%		Homework 7.txt	87%
24	very charged emotively. It's like that I'm <b>overwhelming</b> by some specific			443	3516%		235%				035%		Homework 7.txt	85%
25	meanings. There is many ways for <b>working</b> (to work?) with our			310	2778%		174%				050%		Homework 7.txt	58%
26	reality. They can be a way for a better <b>understanding</b> of our very complex			514	3937%		378%				039%		Homework 7.txt	99%
27	know what to do or think or so on. I'm <b>staying</b> fascinated by the ways choose			477	3615%		278%				032%		Homework 7.txt	91%
28	. His work was really inappropriate for <b>raising</b> girls, so he decided to abandon			212	1135%		047%				047%		Homework 5.txt	49%
29	time and very talented. Strangely, at the <b>beginning</b> of her twenties when she was			76	419%		017%				017%		Homework 5.txt	19%
30	women of the old car 'can that was <b>confining</b> their body. She had a very			56	234%		012%				012%		Homework 5.txt	14%
31	minimum of education and also a basic <b>training</b> for work. Coco Chanel was a			272	1236%		030%				050%		Homework 5.txt	61%
32	Her research for sobriety style has been <b>becoming</b> her famous brand; finish			367	1856%		031%				031%		Homework 5.txt	81%
33	see a very modern view of what she was <b>wanting</b> to be. She built up an empire			345	1633%		076%				076%		Homework 5.txt	76%
34	or future powerful nations are <b>looking</b> for clients, new markets or			172	831%		039%				059%		Homework 4.txt	72%
35	some difficult events. However, when I'm <b>looking</b> at that call our modern world, I			98	421%		039%				039%		Homework 4.txt	41%
36	are looking for clients, new markets or <b>taking</b> their place among the G8 or G21			178	849%		071%				071%		Homework 4.txt	74%
37	the real issue. In fact it depends on how <b>doing</b> it and often against what or who!			228	1053%		031%				031%		Homework 4.txt	91%
38	little dragons that want their place; <b>Taking</b> our place it's not the real issue.			214	973%		036%				035%		Homework 4.txt	87%
39	last swat with its tail of the dragon, I'm <b>hoping</b> for February 9th. I don't really			69	235%		028%				027%		Homework 4.txt	30%
40	have been fascinating the human <b>being</b> ; especially in China, that have			11	042%		0 4%				0 4%		Homework 4.txt	8%
41	For centuries the dragons have been <b>fascinating</b> the human being; especially			8	031%		0 3%				0 3%		Homework 4.txt	7%
42	being; especially in China, that have <b>being</b> occupied a big place in their			17	055%		0 7%				0 7%		Homework 4.txt	11%
43	events have been happening at the <b>beginning</b> of 2013. I said that is			51	133%		020%				020%		Homework 4.txt	24%
44				10	432%		040%				040%		Homework 4.txt	32%

## RESULTS & DISCUSSIONS

### Progressive Aspects & State Verbs

The learner actively uses present and past progressive, and while she struggles with present perfect simple and progressive tenses, she nevertheless attempts to use them whenever possible. She generally has no issues with present and past progressive, as demonstrated in (1) and (2):

- (1) I'm **looking** for a designer which is familiar
- (2) I was **trying** to find my way toward my family

However, she shows a tendency of overusing the progressive instead of simple forms. For example, “I was **waiting** for five years (3)”. In the context of (3), where she tried to recount

something that was over in the past, she should have used “I waited” instead of “was waiting”. This is a common Francophone learner error, as identified by Walter (2001), and consequently an area that requires constant monitoring and scaffolding in her future lessons.

With present perfect progressive, however, mistakes arise sporadically as she continues to improve her command of this grammar point. Sometimes she uses it correctly (4), but sometimes incorrectly (5):

- (4) we have been **searching** solutions,
- (5) sobriety styles has been **becoming** or famous brand.

Example (5) also coincides with her confusion with state verbs. State verbs describe states or conditions and are used only in simple forms, yet learners, including my student, often make mistakes by using state verbs in progressive forms (Parrott, 2010). My student’s mistakes include the following selected examples:

- (6) my mother who was **looking** younger
- (7) I’m **staying** fascinated by the ways choose
- (8) Mother who were **wanting** better life for theirs girls
- (9) theirs feet were **becoming** completely “diform”.
- (10) April is **being** the first with chocolate Easter
- (11) I’m always **remembering** many years ago

In examples (6) to (11), she falsely used various state verbs in present and past progressive aspects. Since this is one of her most common mistakes, it demonstrates a strong awareness of the progressive verb forms yet only an incomplete understanding of its appropriate application. The topic of state verb usage, therefore, is clearly in need of reinforcement.

### **Passive vs. Active Voice**

The learner has frequently demonstrated confusion with passive and active voice in her speech and writing. One of her typical mistakes with voice is illustrated in (12), where she used “overwhelming” instead of “overwhelmed”:

- (12) It’s like that I’m **overwhelming** by some specific

Another example, where she attempted to use passive voice with present perfect progressive: (13) “I have always been **interesting** to the cooking world”, demonstrated a similar issue. In addition, our subsequent discussion as teacher and student (13) revealed her fundamental lack of understanding of the role of past participle in passive voice, as well as the difference between present and past participles because she expressed strong doubt about using “interested” rather than “interesting”. Since these passive examples are also adjectival (Carter & McCarthy, 2006), the use of present and past participles as adjectives should be incorporated and reinforced in future lessons too.

In (14), although she appeared to have constructed a formally correct sentence, she actually intended to write “have fascinated the human being”:

(14) dragons have been **fascinating** the human being

This type of mistake may demonstrate confusion with both tense and voice. In yet another situation, her confusion with state verbs is possibly mixed with voice choice, leading her to create (15):

(15) that have **being** occupied a big place in their

The analysis of her usage of passive and active voice suggests for a pedagogy that should begin by strengthening her understanding between present and past participles, and then gradually building up to passive construction in different tenses, along with a continuing reinforcement of state verb usage.

### Gerunds

Gerunds are when *-ing* forms function as nouns (Parrott, 2010), and this learner usually uses gerunds correctly, as illustrated in (16) to (18):

(16) new tendency of **living** in small apartments

(17) female friends. **Having** male friend can be a problem

(18) I’m afraid of **quitting** what I know very well

However, from my long-term observation of her learning, it is apparent that her successful usage of gerunds and infinitives is not entirely consistent, and she has a tendency of overusing gerunds. It may be speculated that in some cases, her correct usage of gerund is a result of guessing. For example, in (19), “love” is a verb that may take gerunds or infinitives with little difference in meaning, but we tend to use gerunds to state facts; as well, gerunds are typically used to list activities (Parrott, 2010). Therefore, even though the node “moving” was used correctly, (19) demonstrated a level of inconsistency in her choice between gerunds and infinitives.

(19) I love to dance and to run and **moving** for me is life.

In examples (20) and (21), however, she makes clear mistakes by using gerunds after verbs that should be followed by infinitives. Making the distinction between verbs that take gerunds or infinitives has been one of her long-term weaknesses, which she generally tries to compensate for by overusing gerunds. This issue remains a teaching focus in her training.

(20) because the market starts **decreasing**. I know that

(21) a lot of people choose **being** on the warm beaches of south

A different type of mistake was observed in this example, “I had a group who **witnessing** me in a studio” (22). Student-teacher discussion revealed that she actually meant to say, “a group to witness me”. It may be presumptuous to conclude whether this mistake was the

confusion between *-ing* forms and infinitives, but her tendency of overusing *-ing* forms and adjective clauses may have led to the result.

Two special cases of gerund usage may be seen in the titles: "Moving" and "Cooking" (line 77 & 78 of Appendix B). In these cases, the learner showed clear knowledge of using gerunds as nouns and applied them as proper nouns, with capitalization, in the titles of her articles (Carter & McCarthy, 2006). Overall, she demonstrates a basic understanding of the concept of gerunds and infinitives but has yet to successfully internalize and apply the concept, making it an important topic for future lessons.

### **Tenseless *-ing* Phrases**

Tenseless verb phrases or clauses with no stated subjects are generally understood to share the subject with the main clauses, and the tenseless phrases or clauses often have the adverbial function of, for example, extending the main clauses by adding extra information (Carter & McCarthy, 2006; Leech *et al.*, 2006). Such phrases can be constructed with *-ing* participles, and my learner demonstrated her ability of using such phrases in examples (23) and (24).

(23) were very well dressed, arrived **bringing** a large tray.

(24) I'll pass the next few days **finishing** me big clean-up of my

These sentences, albeit correct, were described by the learner as more "accidental". She may not have consistent command of such structures yet, but it would be interesting to continue observing her progress in this direction.

### **Adjectives**

The *-ing* verb forms may function as adjectives and a possible learner mistake is mixing present and past participles in this regard (Carter & McCarthy, 2006; Parrott, 2010). My learner showed a solid command of using *-ing* form adjectives, as shown in examples (25) to (27); (26) is a combination of concordance line 72 & 74 (Appendix B).

(25) It will be very **interesting** to create this new

(26) for its **nourishing, fortifying**, and aphrodisiac virtues

(27) can be very **depressing!**

Moreover, she knows how to create her own adjective: to express my "**mothering**" energy (28). This example shows her understanding of creating adjectives with suffixes (Parrott, 2010). However, it is unclear why she chose the suffix *-ing* instead of, for example, *-ly* for "motherly", or simply the adjective "maternal". Regardless of correctness, this example shows a level of creativity in her English usage. Despite using adjectives correctly in active sentences, she appears to be less consistent with adjectival passive constructions as discussed earlier – an important topic for future lessons.



## CONCLUDING REMARKS & APPLICATION TO CLASSROOM PRACTICE

The frequency of grammatical errors provides constructive implications for material choice and teaching direction (Biber & Reppen, 2002). The main goal of this analysis was to identify the patterns and frequency of learner language in terms of grammar and context and to apply these findings to plan suitable pedagogy and improve future lessons for the learner. The results demonstrated that the learner's weakest points were 1) overusing gerunds and present participles with state verbs; 2) falsely using progressive instead of simple verb forms and 3) confusing passive and active voice, which led to false usage of present participles instead of past participles, or vice versa. The positive aspects of her language were the sound command of adjectives and capability of using tenseless phrases to add extra information. The learner's current level of proficiency no longer requires basic grammar lessons. Instead, a continuing push for production with targeted accuracy refinement could be the more efficient approach.

While a learner corpus clearly and quickly displays learner mistakes, it is nevertheless difficult to assess, on paper, the degree of significance these mistakes represent. As all erroneous sentences have been corrected and discussed with the learner during our lessons, I could observe how she reacts to different mistakes in real time. Some mistakes might have been the result of carelessness and not a lack of knowledge. Through the discussions I, as the instructor, have become highly aware of the true nature of her learning obstacles. As learner-awareness in pedagogical materials is key in the applications of learner corpora (Granger, 2003), the insights gained through classroom discussions of her mistakes would undoubtedly help further refine my lesson planning.

In a one-to-one tutorial context, a combination of corpus analysis and real-life discussion appears to bring the most realistic representation of learner needs and, therefore, would be ideal for the development of future lesson planning. For instance, the same practice can be expanded and adapted for a larger group of students in classroom settings. Other than writing or literacy classes, some other English proficiency courses at Kanda University of International Studies (KUIS) (e.g. Freshman English) ask students to write study reflections regularly. If these reflections are written in electronic texts (e.g. Google Docs), then they can be collected and analyzed using the methods outlined in this paper. The analysis can be done for individual learners, small groups in a class, or for an entire class. Regardless of sample size, a learner corpus analysis will identify clear patterns and frequency of learner language and provide guidance for teachers to develop lessons based on the learners' needs. In addition, learners can also take their corpus analysis results to learning advisors and ask for more specific and direct assistance on the problem area. The application of learner corpora is vast and I highly encourage all language educators to explore this method on their journey to better understand their learners.

## REFERENCES

Biber, D., & Reppen, R. (2002), What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition*, 24, 199-208.

Carter, R., & McCarthy, M. (2006). *The Cambridge grammar of English*. Cambridge: Cambridge University Press.

Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. London: Routledge.

Flowerdew, L. (2012). *Corpora and language education*. Basingstoke: Palgrave Macmillan. Retrieved from <https://www.dawsonera.com/readonline/9780230355569>

Granger, S. (2003). The international corpus of learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly*, 37(3), 538-546.

Leech, G., Deucher, M., & Hoogenraad, R. (2006). *English grammar for today: A new introduction* (2<sup>nd</sup> ed.). Palgrave Macmillan.

McEney, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies*. London: Routledge.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press. Retrieved from <https://www.dawsonera.com/readonline/9780511321733>

Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge: Cambridge University Press.

Scott, M., (2013). *WordSmith tools help*. Liverpool: Lexical Analysis Software.

Walter, C. (2001). *French speakers*. In M. Swan, & B. Smith, (Eds.), *Learner English: A teacher's guide to interference and other problems* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press, pp. 52-72.

## APPENDIX A. Sample material from the corpus

<header>HOMEWORK 9: 6 mars 2013</header>

### SPRING

I know that is not still the spring but I could see and feel some signs of it. The first thing that I notice is the light during the day. In fact, the sun arises earlier and goes down later. What I like is the warm and the irradiance of this light, even if we are in winter; a real antidote against the end of the winter that often can seem long for some people. In Quebec, it seems that is the moment during the year that the sun is physically the closest of us. I love these days which invite us being outside practicing Winter sports with friends and family. Fortunately at this same time, we have these days break for many students and parents; many of them plan trips and outdoors in family, couple or with friends. There are many cultural activities that we can do in town. However, when the weather is too grey, a lot people choose being on the warm beaches of south; there are many travelling packages especially for this period of time.

This year there are no beaches or any trip organised for me. In fact, I have to prepare my move toward a new «nest». I'll pass the next few days finishing me big clean-up of my actual condo. I live here for almost thirteen years now; last year I bought a new condo not too far from this one. I have some butterflies in the stomach because this change is a bit risky. My actual situation at work is unpredictable and my new place is just in front of it. Also the new one is smaller, than half. I'm afraid that will be too small.

I try to figure out living in this «bubble». I know that I can live with few furniture, in a very Zen space. Last month, I watched a documentary on this new tendency of living in small apartments in Vancouver, where the real estate market is very high. People talked about their appreciation of this life style. I could understand them because the price of the houses is so high. This is not the same situation for the Montrealer Market but, in my concern, I want only a «pied à terre» and finding something else in the country.

I try to imagine this small place. I have to sale my old and big furniture. I visualize how to create a comfortable and warm nest for myself. I really like the Japanese style with the doors which slide. I'm looking for a designer which is familiar with this style. I have to find smart ways about storing things. This is my challenge for next weeks and, despite my fears, I'm very excited to do it. Now I have to sale my actual condo and rapidly, because the market starts decreasing. I know that there are many things to do and think; fortunately the arrival of spring will give me a great bust of energy.

## APPENDIX B. The concordance of present participle usage in the corpus

1 A documentary on this new tendency of **living** in small apartments in Vancouver  
2 I want only a "pied a terre" and **finding** something else in the country. I  
3 I'll pass the next few days **finishing** me big clean-up of my actual  
4 be too small. I try to figure out **living** in this "bubble". I know that I  
5 I have to find smart ways about **storing** things. This is my challenge for  
6 rapidly, because the market starts **decreasing**. I know that there are many  
7 style with the doors which slide. I'm **looking** for a designer which is familiar  
8 warm beaches of south; there are many **travelling** packages especially for this  
9 days which invite us being outside **practicing** Winter sports with friends and  
10 is too grey, a lot of people choose **being** on the warm beaches of south;  
11 us. I love these days which invite us **being** outside practicing Winter sports.  
12 for me is life. Ten years ago, I was "moving" with some people; dancers,  
13 body. I love to dance and to run and **moving** for me is life. Ten years ago, I  
14 not as well that if I had a group who **witnessing** me in a studio. However, this  
15 still was on the campus McGill; I was **trying** to find my way toward my family  
16 was younger that in reality, she was **preparing** a dessert for me. There also  
17 were very well dressed, arrived **bringing** a large tray. It's at this time  
18 saw a horse, without cavalier, it was **bringing** a saddle, an English saddle.  
19 and it lay down close to me, it was **protecting** me from the bad weather.  
20 Christmas and my mother who was **looking** younger? I really don't know.  
21 by some specific concerns and I'm **cutting** off them for many reasons: I  
22 emotionally. It's like that I'm **overwhelming** by some specific  
23 meanings. There is many ways for **working** (to work?) with our  
24 what to do or think or so on. I'm **staying** fascinated by the ways choose  
25 His work was really inappropriate for **raising** girls, so he decided to abandon  
26 women of the old car can that was **confining** their body. She had a very  
27 research for sobriety styles has been **becoming** or famous brand; finish  
28 a very modern view of what she was **wanting** to be. She built up an empire  
29 or future powerful nations are **looking** for clients, new markets or  
30 difficult events. However, when I'm **looking** at that call our modern world, I  
31 looking for clients, new markets or **taking** their place among the G8 or G21  
32 real issue. In fact it depends on how **doing** it and often against what or who!  
33 little dragons that want their place; **Taking** our place it's not the real issue.  
34 swat with its tail of the dragon, I'm **hoping** for February 9<sup>th</sup>. I don't really  
35 For centuries the dragons have been **fascinating** the human being; especially  
36 being; especially in China, that have **being** occupied a big place in their  
37 Liette because many events have been **shocking** and we have been searching  
38 have been shocking and we have been **searching** solutions. I have been  
39 been searching solutions. I have been **passing** through different feelings:  
40 For twelve years I have been **working** at the Centre  
41 university campus and everybody are **looking** for financial support. We don't  
42 has to search and to find solutions **keeping** the center alive; the coach that  
43 That's for sure, I'm afraid of **quitting** what I know very well and what I  
44 I try to stay calm and present. I'm **observing** movements around me; I'm  
45 observing movements around me; I'm **talking** with people concerned by the  
46 for the child, if something has to be **happening**. I find that to be a  
47 I feel like a fairy. It will be very **interesting** to create this new  
48 will take still too much time. I was **waiting** for five years. This decision was  
49 There is one year ago (already) I was **deciding** to stop my adoption project in  
50 but I will be possible to express my "mothering" energy! So I really wait for  
51 area. Another thing that I like about **cooking** is the movies about.  
52 is brought great satisfactions like **receiving** family and friends, to create  
53 because I can share my passion of **cooking** with them. Sometime,  
54 I can remember, I have always been **interesting** to the cooking world. I love  
55 from and how we have to prepare them. **Cooking** is more than that to prepare a  
56 I have always been interesting to the **cooking** world. I love to work with the  
57 We are all connected. Also, **cooking** is brought great satisfactions

58 very pertinent proverb here about **feeding**; "when you eat something, you  
59 friendship was for life. They were **learning** a secret language that they  
60 of friendship. Sometime, a mother was **asking** for a friend for her young girl,  
61 very close between them. They were **developing** a deep sens of friendship.  
62 a secret language that they were **using** all their life, to be in touch together  
63 I have male and female friends. **Having** male friend can be a problem,  
64 their life, to be in touch together. **Watching** this movie, reminds me my  
65 a better life for theirs girls, were **asking** to the person who was  
66 feet's women. Mother who were **wanting** better life for theirs girls, were  
67 time, on the 19<sup>th</sup> century, we were **banding** feet's women. Mother who  
68 young girls; theirs feet were **becoming** completely "difform". More  
69 roles were having babies and were **taking** care about them, their family and  
70 of these deals. Women's roles were **having** babies and were taking care  
71 feet were "beautiful", more they were **having** chance to do a great marriage.  
72 They appreciate it for its **nourishing**, fortifying and aphrodisiac  
73 his or her love this day. I'm always **remembering** many years ago; I had a  
74 appreciated it for its nourishing, **fortifying**, and aphrodisiac virtues. Also  
75 report income tax and et can be very **depressing**!  
76 of the year, April is **being** the first with chocolate Easter:  
77 **Moving** In my childhood, I always loved  
78 **Cooking** I don't know why, since I can